Lin_Alg_Glossary.tex compiled April 21, 2006
Tyler Neylon

Linear Algebra

a brief glossary

An $m \times n$ matrix A has m rows and n columns. An $n \times 1$ matrix x consisting of one column is a **column vector**, and writen $x \in \mathbb{R}^n$. If a_i is the i^{th} column of A, then we may write the product Ax as

$$Ax = \sum_{i=1}^{n} a_i x_i,$$

where x_i is the i^{th} coordinate of x.

If B is an $n \times p$ matrix, and $C = AB$, then we may write c_i, the i^{th} column of C as

$$c_i = Ab_i,$$

where b_i is the i^{th} column of B.

The **column-row expansion** form for $C = AB$ is

$$C = \sum_{i=1}^{n} a_i b^i,$$

where b^i is the i^{th} *row* of B. Note that each term $a_i b^i$ here is an $m \times p$ matrix, not to be confused with $a^i b_i$, which is just a scalar (here a^i is the i^{th} row of A).

By $\mathsf{col}(A)$, called the **column space** of A, we indicate the vector space spanned by the columns of matrix A. Specifically,

$$\mathsf{col}(A) = \{Ax : x \in \mathbb{R}^n\}.$$

Similarly, $\mathsf{row}(A)$ indicates the **row space** of A:

$$\mathsf{row}(A) = \{yA : y \text{ is a row vector in } \mathbb{R}^m\}.$$

A **row vector** is a $1 \times m$ matrix (any matrix with only one row).

The **null space** of a matrix A, written $\mathsf{null}(A)$, is the vector space given by

$$\mathsf{null}(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

The columns of A are **linearly independent** iff $\mathsf{null}(A)$ contains only the zero vector.

The **rank** of a matrix A, written $\mathsf{rank}(A)$, is the dimension of $\mathsf{col}(A)$, defined rigorously as the size of the largest set of linearly independent vectors in $\mathsf{col}(A)$. It is a fact that

$$\dim(\mathsf{col}(A)) = \dim(\mathsf{row}(A))$$

for all matrices A. A matrix has **full rank** when $\mathsf{rank}(A) = \min(m, n)$ (this makes sense since necessarily $\mathsf{rank}(A) \leq \min(m, n)$.

The **corank** of a matrix A is the dimension of $\mathsf{null}(A)$. It is a fact that, for all $m \times n$ matrices A,

$$\mathsf{rank}(A) + \mathsf{corank}(A) = n.$$

For some matrix A, it is traditional to write a_{ij} to indicate the value in the i^{th} row and j^{th} column. A matrix is **diagonal** iff $(i \neq j \implies a_{ij} = 0)$; that is, all entries not on the diagonal are zero. A matrix is **upper-triangular** iff $(i > j \implies a_{ij} = 0)$.

Given a matrix A, we say that $B = A^T$ is the **transpose** of A iff $b_{ij} = a_{ji}$. We say that A is **lower-triangular** iff A^T is upper-triangular.

The $n \times n$ **identity matrix**, usually denoted with the particular letter I, has entry I_{ij} in row i and column j given by

$$I_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

It is also traditional to write e_i for the i^{th} column of this matrix.

An $m \times n$ matrix A is **invertible** iff it is square $(m = n)$ and there exists another matrix, written A^{-1} such that $AA^{-1} = A^{-1}A = I$. It is a fact that A is invertible iff its columns are linearly independent iff $\mathsf{rank}(A) = n$ iff $\mathsf{null}(A) = \{0\}$.

The columns of a matrix Q are called **orthonormal** iff

$$q_i^T q_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise,} \end{cases}$$

where q_i is the i^{th} column of Q. A real square matrix Q is **unitary** iff its columns are orthonormal. It is a fact that a matrix is unitary iff $Q^T = Q^{-1}$ iff its rows are orthonormal.

Matrix U is in **reduced row echelon** form iff

- there is an increasing sequence $P = \{p_1, \ldots, p_r\} \subset [n]$ of column indices so that column $u_{p_i} = e_i$, the i^{th} column of the identity matrix; and
- for $i \leq r$, $j < p_i \implies u_{ij} = 0$; and
- for all $i > r$, $u_{ij} = 0$,

where $r = \mathsf{rank}(U)$. The set of columns indexed by P are the **pivot columns** in U. This form is what we reduce a matrix to by Guassian elimination. Notice that any U in reduced row echelon form is also upper-triangular.

Reduced row echelon example An *elementary row operation* on a matrix A is a simple operation such as

- switching two rows,

- multiplying a row by a scalar, or
- replacing a row r by $r + s$, where s is another row in the matrix.

It is a fact that matrix B may be derived from matrix A by elementary row operations iff there is an invertible matrix X such that $B = XA$. Thus we may always algebraically represent a series of row operations by some matix X which is to be left multiplied by A (that is, our result is given by XA).

As a quick example, consider

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 4 \\ 3 & 6 & 5 \end{pmatrix}.$$

We can use the upper-left entry as a pivot in Guassian elimination to arrive at

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

This matrix is upper-triangular but not yet in row echelon form — although the first column qualifies as a pivot column, the third does not, and there is no assignment of pivot columns which would allow this matrix to meet the conditions of being in reduced row echelon form. Intuitively, the problem is that the third column "goes down by two" at a time, which is not allowed in row echelon form.

So our next step will be to exercise another row operation in order to clean up the third column:

$$C = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Technically, this is still not in reduced row echelon form since the third column is still not of the form e_i, although we could easily solve a system of the form $Cx = b$ at this point.

Let's eliminate the upper-right nonzero, and then divide the third column by 2 in order to arrive at:

$$D = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

At last, we have reached reduced row echelon form. Indeed, let $p_1 = 1$ and $p_2 = 3$ be our pivot columns. No column "goes down" by more than one row at a time, and those columns which do (the pivot columns) are identity columns (from an identity matrix).

Decompositions The **QR decomposition** of $m \times n$ matrix A is

$$A = QR,$$

where Q is $m \times m$ and unitary, and R is upper triangular. Every matrix has a QR decomposition.

The **LU decomposition** of $m \times n$ matrix A is

$$A = LU,$$

where L is invertible and U is in reduced row echelon form. As we have defined it, every matrix has an LU decomposition. (Many authors prefer to require that L is also lower-triangular, in which case this decomposition may not exist, although it will when one allows a row permutation to take place. This is where the term 'psycholocially row echelon form' comes from, meaning 'up to a row permutation.' I plan to eliminate this term from the thesis so it will not be necessary.)

The **singular value decomposition** (SVD) of $m \times n$ matrix A is

$$A = U \Sigma V^T,$$

where U is $m \times m$ and unitary, V is $n \times n$ and unitary, and Σ is $m \times n$ and diagonal. In addition, if σ_i indicates the i^{th} value along the diagonal of Σ, we require that $\sigma_1 \geq \sigma_2 \geq \ldots$ and that each σ_i be nonnegative. All real matrices have a real singular value decomposition.

Norms Given two column vectors x and y, their **inner product**, sometimes written as $\langle x, y \rangle$, or just $x \cdot y$, is simply the matrix product

$$\langle x, y \rangle = x^T y.$$

A **norm** is a function mapping $\mathbb{R}^n \to \mathbb{R}$ which is nonnegative, zero only when input x is zero, scales linearly with the input, and obeys the triangle inequality. In notation:

- $||x|| \geq 0$ and $||x|| = 0$ iff $x = 0$;
- $||\lambda x|| = |\lambda| \cdot ||x||$ for any scalar λ; and
- $||x + y|| \leq ||x|| + ||y||$.

The standard norm, which we denote simply by $||x||$ is defined as

$$||x|| = \left(\sum_i x_i^2 \right)^{1/2}.$$

This may also be written as $||x||_2$ in cases where the context may suggest otherwise.

The **Cauchy-Schwarz** inequality tells us that

$$\langle x, y \rangle \leq ||x|| \cdot ||y||$$

for any pair of vectors x, y.

Another norm we will use occasionally is given by

$$||x||_1 = \sum_i |x_i|.$$